BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PATENT

Attorney Docket No. 080135-000000

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:

CHRISTOPHER J. ONG, *et al.*

Serial No.: 09/295,464

Filed: April 19, 1999

For: COMPLEMENTATION TRAP

Examiner: R. Schnizer

Art Unit: 1632

TRANSMITTAL OF PRIORITY DOCUMENT

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

In conjunction with the above-referenced application, Applicants submit herewith a certified copy of Canadian priority Application No. 2,205,888, filed February 11, 1997, thereby perfecting Applicants' priority claim.

Respectfully submitted,

Eugenia Garrett-Wackowski

Eugenia Garrett-Wackowski
Reg. No. 37,330

TOWNSEND and TOWNSEND and CREW LLP
Two Embarcadero Center, 8th Floor
San Francisco, California 94111-3834
(415) 576-0200, Fax (415) 576-0300
EGW/lls

Office de la propriété intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An Agency of Industry Canada

Bureau canadien des brevets

Certification

Canadian Patent Office

Certification

La présente atteste que les documents ci-joints, dont la liste figure ci-dessous, sont des copies authentiques des documents déposés au Bureau des brevets.

This is to certify that the documents attached hereto and identified below are true copies of the documents on file in the Patent Office.

Specification and Drawings, as originally filed, with Application for Patent Serial No: **2,205,888** on July 11, 1997, by **THE UNIVERSITY OF BRITISH COLUMBIA,** assignee of Christopher J. Ong, John J. Priatel and Frank R. Jirik, for "Complementation Trap".

Agent certificateur/Certifying Officer

November 6, 2000

Date

Canada

(CIPO 68)

OPIC CIPO

COMPLEMENTATION TRAP

Field of the Invention

The present invention relates to entrapment vectors and their use in gene discovery, and their use in screening for or making cells and organisms that are mutated for such genes. This invention also relates to the use of entrapment vectors to identify tissue specific transcription control elements such as promoters and enhancers and for generating transgenic animals displaying restricted expression of transgenes. This invention also relates to trap vectors comprising a splice acceptor, a sequence encoding β-galactosidase and an optional IRES sequence.

Background of the Invention

Progress of genomic based drug discovery is largely dependent upon the identification of genomic targets. Thus, cloning, sequencing, and identification of function of mammalian genes is a priority. In particular, it is important to identify and make use of genes which are spacially and/or temporally regulated in the organism.

Animal model systems such as the fruit fly and the worm are often used in gene identification because of ease of manipulation and ability screen for mutants. While these systems have their limitations, large numbers of developmental mutations have been identified in those organisms either by monitoring the phenotypic effects of mutations or by screening for expression of reporter genes incorporated into developmentally regulated genes.

Many features of the mouse make it the best animal model system to study gene function. However, the mouse has not been used for large scale classical genetic mutational analysis because

random mutational screening and analysis is very cumbersome and expensive due to long generation times and maintenance costs.

A disadvantage in using animal models for the identification of genes is the need to establish a transgenic animal line for each mutational event. This disadvantage is alleviated in part by using embryonic stem (ES) cell lines because mutuational events may be screened *in vitro* prior to generating an animal. ES cells are totipotent cells isolated from the inner cell mass of the blastocyst stage embryo. Methods are well known for obtaining ES cells, incorporating genetic material into ES cells, and promotion of differentiation of ES cells. ES cells may be caused to differentiate *in vitro* or the cells may be incorporated into a developing blastocyst and thereby contribute to all differentiated tissues of the resulting animal. Vectors for transforming ES cells and suitable genes for use as reporters and selectors are also well known.

Entrapment vector strategies have been employed to identify developmentally regulated genes. One such vector called a "promoter trap" consists of a reporter gene sequence lacking a promoter. Its integration is detected when the reporter is integrated "in-frame" into an exon. "Gene trap vectors" target the much more prevalent introns of the eucaryotic genome. The latter vectors consist of a splice-acceptor site upstream from a reporter gene. Integration of the reporter into an intron results in a fusion transcript containing from the endogenous gene and the reporter gene sequence.

Gene trap vectors may be made more efficient by incorporation of an internal ribosomal entry site (IRES) such as that derived from the 5′ non-translated region of encephalomyocarditis virus (EMCV). Placement of a IRES site between the splice acceptor and

the reporter gene of a gene trap vector means that reporter gene product need not be translated as a fusion product with the endogenous gene product, thereby increasing the likelihood that integration of the vector will result in expression of the reporter gene product.

Examples from the literature of the use of promoter and gene trap vectors as well as such vectors including an IRES sequence, are listed below. Some examples involve the identification of developmentally regulated or tissue specific events making use of ES cell lines.

1. Canadian Patent application no. 2,166,850 (open for public inspection July 11, 1996) Vectors and the Use Thereof for Capturing Target Genes: describes the use of transmembrane sequence encoding gene trap vectors to isolate and identify secretory proteins.

2. US Patent 5,364,783 issued Nov. 15, 1994. Retrovirus Promoter Trap Vectors: describes retroviral vectors that are used to isolate transcriptionally active chromosomal regions and to identify promoter sequences. The reporter gene is placed in the U3 or U5 control region of the retrovirus.

2. Gossler, A., et al. (1989). Mouse Embryonic Stem Cells and Reporter Constructs to Detect Developmentally Regulated Genes. Science 244:463-465: describes the use of enhancer trap gene trap vectors for use in identifying developmentally regulated genes. The gene trap vector consists of the mouse En-2 splice acceptor upstream from lacZ (reporter) and a selector gene (hBa-neo).

3. Von Melchner, H., et al.: Isolation of Cellular Promoters by Using a Retrovirus Promoter Trap. Proc. Natl. Acad. Sci. USA 1990, 87:3733-3737.

3. MacLeod, D., et al.: A Promoter Trap in Embryonic Stem (ES Cells Selects for Integration of DNA into CpG Islands. Nucleic Acids Res. 1991, 19:17-23.

4. Reddy, S., et al.: Retrovirus Promoter-Trap Vector to Induce lacA Gene Fusions in Mammalian Cells. J. Virol. 1991, 65:1507-1515.

5. Brenner, D.G., et al.: Analysis of Mammalian Cell Genetic Regulation in Situ by Using Retrovirus-Derived Portable Exons Carrying the Escherichia coli lacZ Gene. Proc. Natl. Acad. Sci. USA. 1989, 86:5517-5521.

6. Kerr, W.G., et al.: Transcriptional Defective Retroviruses Containing lacZ for the in Situ Detection of Endogenous Genes and Developmentally Regulated Chromatin. Cold. Spring. Harb. Symp. Quant. Biol. 1989, 54:767-776.

7. Friedrich, G. and Soriano, P.: Promoter Traps in Embryonic Stem Cells: A Genetic Screen to Identify and Mutate Developmental Genes in Mice. Genes. Dev. 1991, 5:1513-1523.

8. Skarnes, W.C., et al.: A Gene Trap Approach in Mouse Embryonic Stem Cells: The lacZ Reporter is Activated by Splicing, Reflects Endogenous Gene Expression, and is Mutagenic in Mice. Genes Dev, 1992, 6:903-918: shows that gene trapping results in activation of lacZ by splicing to endogenous exons and expression as a fusion protein whose expression pattern mimics that of the endogenous gene. The resulting integration and fusion is mutagenic. Using 5' RACE, the endogenous gene activated with three lacZ-based gene-trap insertions was cloned and the proper use of the En-2 splice acceptor site was demonstrated. For two insertions, the pattern of lacZ expression in embryos was shown to

match the normal distribution of endogenous transcripts. Two of the three insertions tested cause phenotypic abnormalities in mice. One of those was an insertion into a novel gene expressed widely during development that causes perinatal death in homozygous animals. The other is an insertion into a zinc-finger gene expressed in neural cells that results in mild growth retardation after birth.

9. Von Melchner, H., et al.: Selective Disruption of Genes Expressed in Totipotent Embryonal Stem Cells. Genes. Dev. 1992, 6:919-927: where sequences upstream of nine retroviral promoter-trap insertions were cloned using inverse PCR. Flanking probes from five ES cell lines detected transcripts, and one clone is identified as the REX-I transcription factor. Two of four lines transmitted to the germline caused embryonic-lethal phenotypes.

10. Sheriden, U., et al.: Transcriptionally Active Genomic Regions are Preferred Targets for Retroviral Integration. Mol. Cell. Biol. 1990, 64:907-912.

11. Vijaya, S., et al.: Acceptor Sites for Retroviral Integrations Map Near DNase 1-Hypersensitive Sites in chromatin. J. Virol. 1986, 60:683-692.

12. Rohdewold, H., et al.: Retrovirus Integration and Chromatin Structure: Moloney Murine Leukemia Proviral Integration Sites Map near DNAse I Hypersensitive Sites. J. Virol 1987, 61:336-343.

13. Boggs, S.S., et al.: Efficient Transformation and Frequent Single Site, Single Copy Transcription of DNA can be Obtained in Mouse Erytholeukemia Cells Transformed by Electroporation. Exp Hematol 1986, 149:988-994.

14. Soininen, R., et al.: The Mouse Enhancer Trap Locus I (Etl-1): A Novel Mammalian Gene Related to Drosophila and Yeast Transcriptional Regulator Genes. Mech Dev 1993, 39:111-123.

15. Niwa, H., et al.: An Efficient Gene-Trap Method using Poly A Trap Vectors and Characterization of Gene-Trap Events. J. Biochem 1993, 113:343-349.

16. Breindl, J.K., et al.: Retrovirus-Induced Lethal Mutation in Collagen I Gene of Mice is Associated with Altered Chromatin Structure. Cell 1984, 38:9-16.

17. Jahner, D. and Jaenisch, R.: Retrovirus-Induced de Novo Methylation of Flanking Host Sequences Correlates with Gene Inactivity. Nature 1985, 315:594-597.

18. Kratochwil, K.: Retroviral-Induced Mutation in Mov 13 Mice Affects Collages I Expression in a Tissue-Specific Manner. Cell 1989, 57:807,816.

19. Goebl, M.G. and Petes, T.D.: Most of the Yeast Sequences are Not Essential for Cell Growth and Division. Cell 1986, 46:983-922.

20. Wilson, C., et al.: P-element Mediated Enhancer Detection: An Efficient Method for Isolating and Characterizing Developmentally Regulated Genes in Drosophila. Genes. Dev. 1989, 3:1301-1313.

21. Wood, S.A., et al.: Non-Injection methods for the Production of Embryonic Stem Cell-Embryo Chimeras. Nature 1993, 365:87-89: describes time-saving methods for generating ES cell derived chimeras by morulae aggregation. These can be used to more easily

carry out a prescreen based on embryonic expression of the reporter.

22. Doetschman, T.C.: The in Vitro Development of Blastocyst-Derived Embryonic Stem Cell lines: Formation of Visceral Yolk Sac, Blood Islands and Myocardium. J. Embryo. Exp. Morph. 1985, 97:27-45.

23. Risau, W., et al.: Vasculogenesis and Angiogenesis in Embryonic-Stem-Cell Derived Embryoid Bodies. Development 1988, 102:471-478.

24. Wiles, M.V. and Keller, G.: Multiple Hematopoetic Lineages Develop from Embryonic Stem (ES) Cells in Culture. Development 1991, 111:259-267.

25. Rossant, J., et al.: A large Scale Gene Trap Screen for Insertional Mutations in Developmentally Regulated Genes in Mice. Genetics 1995, 139:889-899: describes the use of the gene trap vector to mutagenize ES cells and the screening of ES clones by injection into blastocysts and assaying for reporter gene expression in 8,5 dpc chimeric mouse embryos.

26. Thompson, et al.: Gene Trapping in Differentiating Cell Lines: Regulation of the Lysosomal Protease Cathepsin B in Skeletal Myoblast Growth and Fusion: describes the use of a retroviral promoter trap vector in a myoblast cell line called C2C12 to identify and isolate skeletal muscle specific genes.

27. Shiria, et al.: A Gene Trap Strategy to Identify Genes That Are Expressed in the Developing Mouse Nervous System. Zoological Science 1996, 13:277-283: describes a gene trap vector called pSneoNlacZA in ES cells. ES clones were allowed to differentiate

in vitro in a suspension culture system into neurons, neuro-glia precursors cells. Neuronal cells were identified by using antibodies reactive against neuron specific proteins. Three clones that were immunohistochemically positive and β-gal positive were identified.

28. Imai, et al.: Cloning of Retinoic Acid Induced Gene, GT1, in Embryonal Carcinoma Cell Line P19: Neuron Specific Expression in Mouse Brain. Molecular Brain Research. 1995, 31:1-9: describes the use of a gene trap vector in mouse P19 embryonal carcinoma cell line which can be reproducibly differentiated to neurons and glial cells upon treatment with retinoic acid. Assay for up or down regulation of β-gal activity and immunohistochemistry showed localization in neurons.

29. Forrester, et al.: An Induction Gene Trap Screen in Embryonic Stem Cells: Identification of Genes that Respond to Retinoic Acid *in vitro*. Proc. Nat. Acad. Sci. USA 1996, 93:1677-82: where a gene trap vector was used in ES cells and treated with retinoic acid, with a screen for genes that are induced or repressed upon retinoic acid exposure.

30. Hill, D.P. and Wurst, W. (1993): Screening for Novel Pattern Formation Genes Using Gene Trap Approaches, Methods in Enzymology 225:664-681.

31. Kim, D.G., et al. (1992): Construction of a Bifunctional mRNA in the Mouse by Using the Internal Ribosomal Entry Site of the Encephalomyocarditis Virus. Molecular and Cellular Biology, 12: 3636-3634: describes construction of vectors comprising a promoter driven selector (eg. neo) followed by EMCV 5′ non-translated region and either lacZ or cat, and transfection of murine ES cells with the vectors. The IRES of the EMCV permitted more efficient

translation of the reporter in resulting chimeric embryos which also expressed the selection gene.

32. Chowdhury, K. et al. (1997): Evidence for the Stochastic Integration of Gene Trap Vectors into the Mouse Germline. Nucleic Acids Research, 25: 1531-1536: describes construction of a gene trap vector comprising an EN-2 splice acceptor followed by an IRES sequence and B geo (lacZ reporter fused with neo selector). Murine ES cells were transfected and neomycin resistant colonies were monitored for β-galactosidase activity indicative of lacZ expression Molecular cloning of trapped exons was carried out using 5'-RACE.

33. D.P. Hill and W. Wurst (1993): Screening for Novel Pattern Formation Genes Using Gene Trap Approaches; Methods in Enzymology 225: 664-681.

34. G. Friedrich and P. Soriano (1993): Insertional Mutagenisis by Retroviruses and Promoter Traps in Embryonic Stem Cells; Methods in Enzymology 225: 681-701.

35. W.C. Skarnes (1993): The Identification of New Genes: Gene Trapping in Transgenic Mice; Current Opinion in Biotechnology 4: 684-689.

Regardless of the model, the process of identifying and manipulating tissue specific genes requires elaborate screening procedures to link a mutation to a particular spacial/temporal scheme whereby the mutation is detected in the relevant tissue. It is therefore desirable to make screening and manipulation of new spacial/temporally regulated genes in all eucaryotic models more efficient.

Summary of Invention

This invention makes use of known genes whose expression is restricted to specific tissue or specialized cells ("restricted expression") to facilitate identification and manipulation of new genes and their associated transcription control elements which have similar patterns of expression.

Accordingly, this invention provides a method of providing an indicator of restricted expression of a target gene in a eucaryotic organism, which comprises the steps of:

(i) transforming a eucaryotic cell by placing a DNA sequence encoding a first indicator component under the control of a promoter having restricted expression;

(ii) transforming the cell of (i) or a cell derived from the cell of step (i), by operably integrating into the genome of the cell, DNA lacking a promoter but which includes a sequence encoding a second indicator component;

(iii) producing tissue or specialized cells from the cell of (ii); and

(iv) monitoring the tissue or specialized cells of (iii) for a detectable indicator resulting from the combination of the first and second indicator components.

This invention also provides a method of obtaining a gene, a part of a gene, transcription control element or other nucleotide sequence, having restricted expression which includes isolating endogenous DNA flanking of the sequence encoding the second indicator component from a cell or progeny of a cell prepared as

described above in which the detectable indicator is found. This invention also provides a method of modifying the gene, transcriptional control element or other nucleotide sequence obtained as described above which includes modification of the endogenous DNA which flanks the coding sequence of the second indicator component.

This invention also provides a method of providing a eucaryotic organism producing a detectable indicator in a specialized cell or tissue of the organism which includes growing a multi-cellular eucaryotic organism from a cell or a progeny of a cell prepared as described above in which the detectable indicator is found. The organism may also express in the specialized cell or tissue, the product of heterologous DNA expressed with DNA flanking of the second indicator component of the detectable indicator.

This invention makes use of a complementation trap whereby an indictor is detected only when more than one component of the indicator is expressed in the same cell. One method of this invention involves stable and operable insertion of a nucleic acid construct which includes one indicator component of the complementation trap under the control of a known promoter having restricted expression into the genome of the cell. This is followed by further transformation of the cell, or a cell derived from the previously transformed cell, by insertion into the cell's genome of a complementation trap vector which includes a DNA which encodes a second indicator component. Insertion of the complementation trap vector into an endogenous gene may result in the transcription of both the endogenous gene and the second indicator component under control of the endogenous promoter and subsequent translation of the second component. If the complementation trap vector integrates into an endogenous gene which is expressed in the same cell or tissue type as the first

indicator component, a indicator resulting from the combination of the first and second indicator components may be detected. Thus, this invention allows one to distinguish target eukaryotic genes which are expressed in the same cell or tissue type as the promoter controlling the first component of the complementation trap system.

Accordingly, this invention also provides a first DNA construct comprising splice acceptor upstream of a sequence encoding a first peptide selected from the group consisting of an alpha peptide and an omega peptide of β-galactosidase (β-gal), wherein the first peptide lacks β-gal activity but is capable of alpha complementation to produce active β-gal. This invention also provides a eucaryotic cell and a non-human, multi-cellular eucaryotic organism having operably and stably incorporated into its genome, the first DNA construct described above. This invention also provides a kit including the first DNA construct and, a second DNA construct comprising a promoter having restricted expression upstream of a sequence encoding a second β-gal peptide selected such that the second peptide is capable of alpha complementation with the first peptide. This invention also provides a eucaryotic cell and a non-human, multi-cellular eucaryotic organism having operably and stably incorporated into its genome the first DNA construct and which also expresses the second DNA construct.

The complementation trap vectors useful for this invention, including the above-described first DNA construct, may also contain additional elements such as an internal ribosome entry sequence (IRES) upstream of the second indicator component (eg. the first peptide) or, sequences that allow for directional integration of a heterologous gene, such as a recognition site for site specific recombination (eg. lox).

This invention may be used to identify tissue or cell type specific genes. For example, a nucleic acid construct containing the coding sequence of a first indicator component under control of a known tissue specific promoter is "seeded" into cells which may be made to differentiate or whose nuclear material may be incorporated into cells which are capable of differentiation. The second component is then introduced into a cell having the "seeded" construct, by means of the complementation trap vector. The resulting cells, or cells derived from those cells, are then allowed to differentiate, for example by addition or withdrawal of a chemical inducer/repressor or allowed to spontaneously differentiate. The cells are then screened for activity of the indicator which will occur in the cell or tissue type in which the promoter is functional.

This invention is particularly useful for screening cell or tissue specific genes or transcriptional control elements such as promoters and enhancers, in any animal from which embryonic stem (ES) cell lines may be obtained. The ES cell is subjected to the complementation trap process described above. The ES cells are then allowed to differentiate *in vitro* and cells are screened for the indicator. ES cells demonstrating expression of the indicator may be introduced into the blastocyst stage embryo to produce an animal which will exhibit the same pattern of expression. Where introduction of the complementation trap vector results in disruption of the target gene, the mice resulting from the mutated ES cells will be useful as "knock-out" organisms for the target gene. Alternatively, location of the trap vector into the endogenous gene will provide means for subsequent disruption of the target gene for production of "knock-out" organisms. The complementation trap may be used as a means for locating, cloning, sequencing, and further mutation of the target gene or promoter or enhancer sequences associated with the target gene.

This invention will make new transcriptional control elements available for use in making new transgenic animals displaying cell, tissue or organelle specific expression of transgenes. Promoters made available by use of this invention may themselves be made to control the first indicator component in the complementation trap of this invention in a "leap-frog" procedure whereby new genes having the same restriction pattern as the promoter or may be located.

Brief Description of Drawings

Figure 1: is a schematic illustrating a DNA construct useful in the complementation trap of this invention comprising the myeloid cell specific promoter CD11b controlling the sequence encoding a β-galactosidase alpha peptide terminated by the human growth hormone poly-adenylation signal. Downstream is a selection cassette which in this case confers resistance to hygromycin driven by the phosphoglycerate kinase promoter.

Figure 2: is a schematic illustrating a DNA construct useful as a complementation trap vector of this invention comprising the mouse En-2 splice acceptor upstream of the coding sequence of the β-galactosidase omega peptide and a selection cassette for neomycin resistance.

Figure 3: is a schematic illustrating a complementation trap vector of this invention in which a IRES sequence is positioned upstream of the β-gal omega sequence shown in Figure 2.

Figure 4: is a schematic illustrating a complementation trap vector of this invention in which a IPES sequence and a lox site

are positioned downstream of the omega peptide coding sequence shown in Figure 2.

Description of Specific Embodiments

The term "eucaryotic organism" in this specification means any multi-cellular eucaryote having cells which are differentiated to specific cell or tissue types and includes plant or animal organisms, including, but not restricted to: Drosophila, nematodes, fish and mammals. The method of this invention may be carried out to transform any eucaryotic cell which is capable of differentiation into tissue or specialized cell types and includes any bi- or multipotent eucaryotic cell as well as a differentiated eucaryotic cell whose nuclear material may be transplanted to a bi- or multipotent cell. An example of the latter situation is where differentiated cells are transformed according to the method of this invention and whose nuclear material is subsequently placed in a enucleated ovum or similar vehicle which is then caused to become bi- or multipotent and thus capable of differentiation. Particularly suitable for this invention are totipotent cells such as ES cells.

The term "restricted expression" in this specification means the restriction of a transcription control element (such as a promoter or an enhancer) or the restriction of expression of a gene, such that the aforesaid function or expression occurs in a particular tissue or cell type in a eucaryotic organism. Thus a gene or transcription control element having "restricted expression" is a tissue or cell specific gene or control element.

The "detectable indicator" of this invention is an event which results from the combination of more than one indicator component.

This invention may be used to isolate, characterize, and modify genes that are differentially expressed. Tissue or cell specific genes may be located by using an appropriate tissue or cell specific promoter to drive the first indicator component of the complementation trap system. An example of this invention used to identify genes that are expressed in myeloid cells of the hemopoietic system, would be to make use a known myeloid cell/tissue specific promoter such as CD11b to drive the expression of the first indicator component of the complementation trap system. Other examples of tissue restricted promoters are:

Neural: Neuron specific RIβ sub unit of cAMP-PK promoter, Tryptophan Hydroxylase promoter, Neural specific enolase promoter, Tyrosine hydorxylase promoter, Tα1 α- tubulin promoter;

Lung: α1 collagen gene promoter, rat clara cell 10 protein PROMOTER, human surfactant protein SP-C promoter, preproendothelin promoter;

Liver: human apolipoprotein E promoter;

Heart: alpha B crystallin promoter, murine alpha myosin heavy chain promoter;

Thymus: lck proximal promoter;

T cells: CD2 promoter, CD4 promoter, CD3 promoter;

B cells: IGEμ, CD19 promoter;

Osteoclast - and osteoblast specific promoter: Mouse pro-α1(I) promoter, Tartrate Resistant Acid Phosphatase (TRAP) promoter, CD11b promoter.

One example of a detectable indicator that may be used in this invention is alpha complementation of β-galactosidase. The β-galactosidase enzyme can be divided into two peptides called alpha and omega. β-galactosidase enzyme activity is present only when both alpha and omega peptides are present in the same cell. The alpha or the omega peptide alone shows no detectable enzymatic activity. β-galactosidase enzyme activity can be detected by using commercially available enzyme substrates such as X-gal™ which turns blue upon hydrolysis by β-galactosidase. Other substrates such as Galactan-plus™ produce light upon enzymatic cleavage allowing for detection using a luminometer while other substrates produce fluorescent products detectable by flow cytometry or fluorescence microscopy.

Alpha complementation of beta-galactosidase will work in eucaryotic cells including yeast and mammalian cells; see: Mohler, W.A. and Blau, H.M. 1996. Gene Expression and Cell Fusion Analyzed by lacZ Complementation in Mammalian Cells. Proc. Natl. Acad. Sci. 93:12423-12427; and Moosman, P. and Rusconi, S. (1996). Alpha Complementation of lacZ in Mammalian Cells. Nucleic Acids Res. 24:1171-1172.

In order to identify genes that are expressed in myeloid cells, one could use the CD11b promoter to drive tissue restricted expression in myeloid cells of a sequence encoding one of the β-gal peptides (eg. the alpha subunit comprising about 85 or more N-terminal amino acids of β-gal). The complementation trap vector would include a splice acceptor sequence, a sequence encoding an omega subunit of β-galactosidase (eg. inactive β-gal peptide lacking a N-terminal portion of β-gal), and optionally, an IRES sequence and a lox site for directed integration of heterologous genes. The trap vector may optionally contain a second reporter system to identify and characterize the endogenous gene expression

pattern. The trap vector may also contain a selectable marker cassette.

A second approach is to introduce the first indicator component to be under control of an endogenous promoter having restricted expression endogenous gene via homologous recombination. For example, in order to obtain expression in early immature B cells, a sequence encoding the first indicator component may be introduced by homologous recombination into the CD 19 gene in mouse ES cell germline such that the expression of the indicator component under the control of the CD 19 gene (eg. see: Rickert, R.C. et al. (1997): B Lymphocyte-specific, Cre-mediated Mutagenisis in Mice; Nucleic Acids Research, 25:1317-1318).

Stable lines of ES cells bearing the tissue restricted promoter driving the alpha component are generated and tested for appropriate tissue restricted expression. Mutations are then introduced in the mouse genome by random integration of the omega complementation trap vector. Since the trap vector has a splice acceptor sequence fused to the omega component, integration of the trap vector into a gene intron results in splicing omega mRNA into the transcript of the endogenous gene and ultimately, expression of a fusion protein comprising of a portion of the endogenous protein and the omega subunit. Where a IRES is placed upstream of the omega component, the omega peptide is translated without being fused to the endogenous protein.

Individual integration/mutational events are propagated and replica plated. One fraction of the cells are allowed to undergo *in vitro* differentiation and are assayed for the indicator. If both ß-gal subunits are expressed in the same cell, the indicator is detected whereas the components expressed individually in cells will not be detected. Thus, this procedure identifies

integration/mutational events into genes that are expressed in the same given tissue cell type that is defined by the known tissue restricted promoter driving the first component.

Target genes into which the gene trap has integrated can be cloned and sequenced, for example, by 5' RACE PCR. Undifferentiated ES cells can be used to generate mice mutated for the respective gene. The endogenous tissue restricted promoter of the target gene can be characterised and heterologous genes can be inserted into the site by homologous recombination or site directed recombination, thus allowing the heterologous gene to be driven by the newly identified endogenous tissue restricted promoter.

A "leap-frog" procedure may be employed whereby the tissue specific promoter of the target gene may become the "known" promoter driving expression of the first indicator component. The example employing β-gal complementation described above may be accomplished regardless of which β-gal peptide is expressed in the target gene or under control of the known tissue specific promoter. Thus, cells when containing both β-gal peptide inserts have been shown to result in restricted expression of both peptides upon differentiation, a stored sample of the original transformed cells may be manipulated to remove or disrupt expression of the peptide whose expression is controlled by the "known" tissue restricted promoter. For example, the DNA which encodes the first indicator component under control of the known promoter may be flanked by lox sequences which permit the excision of the sequence encoding the first indicator from the genome with Cre recombinase. A complementation trap vector may then be introduced comprising a splice acceptor and a coding sequence equivalent to that which has been removed or disrupted. Mutational events which result in expression of the new trap vector are detected because of the combination of the indicator component expressed as a result of the

first introduced trap vector, plus the component expressed as a result of the second trap vector. Thus, genes having a pattern of expression similar to the first target gene may be located even if the pattern of expression differs from that of the original "known" promoter.

The foregoing description illustrates an embodiment of this invention where the first component is driven by a known promoter and the detectable indicator results from enzymatic complementation of the expression products of the first and second indicator component coding sequences. This invention may be carried out using any detectable gene product which exhibits intra-cistronic complementation. The following are further examples of different indicators and components thereof which may be used in this invention:

(a) Transcriptional activation (two hybrid system): the transcription of a reporter gene requires the presence of two genetic components; a DNA binding subunit and a transcriptional activating subunit. Both components need to be present in the same cell in order for expression of a reporter gene is achieved as the indicator.

(b) Transcriptional activation (genetic recombination dependent): transcription and expression of a reporter gene regulated by a tissue specific promoter is conditional upon genetic recombination that is controlled by the expression of a recombinase enzyme. For example, a stuffer DNA sequence (flanked by lox sequences) which prevents efficient transcription or expression of a reporter gene is removed by recombination that is mediated by a Cre recombinase enzyme allowing for expression of the reporter gene. Alternatively, a recombination event may be required to juxtapose the promoter/enhancer sequence with the reporter gene in

the correct orientation to allow for the efficient expression of the reporter gene.

(c) Enzyme pathway: the indicator is a detectable product that is generated by the sequential action of two enzymes on a substrate. Alternatively, the products produced by the action of two enzymes interact in such a way as to generate a detectable signal. Alternatively, the activity of a first enzyme is regulated by the activity of a second enzyme, (for example: enzymatic activity of first enzyme is dependant on a post translational modification mediated by a second enzyme such as proteolytic activation, glycosylation or phosphorylation etc.).

(d) Selection/screening system: a selectable marker is driven by a tissue specific promoter and a second reporter gene is used in a trap vector. The cells are induced to differentiate and are subsequently placed under selection that allows for the survival of only those cells expressing the selectable marker. Those cells are then screened for the expression of the second reporter.

(e) Co-expression of two reporter genes: one reporter is driven by a "known" tissue specific promoter and the second reporter is used in a trap vector. Differentiated cells are screened for expression of both reporter genes in the same cell.

Examples of suitable reporter genes that may be employed in the above-described examples are: β-galactosidase, alkaline phosphatase, blue fluorescent protein, green fluorescent protein (GFP), β-glucuronidase (GUS), Cre recombinase, Flp recombinase, luciferase, and their cell surface protein with antibodies that can detect its expression. Suitable selectable markers that may be used are: Neo- (neomycin phosphotransferase), Hygro- (hygromycin

β-phosphotransferase), Puro- (puromycin n-acetyltransferase *pac*), HisD- (histidinol dehydrogenase), and Gpt- (xanthine/guanine phosphoribosyl transferase).

Examples of ES cell lines which may be used in this invention are: porcine (eg. US Patent 5523226 Transgenic Swine Compositions and Methods); murine (eg. D3, R1, CGR8, AB1 ES cell lines); primate (eg. rhesus monkey); rodent; marmoset; avian (eg. chicken); bovine; rabbit; sheep; and horse.

A wide variety of cells may be targeted by the trap vectors of this invention, including stem cells, ploripotent cells such as zygotes, embryos, ES cells, other cells such as lymphoid and myeloid stem cells, neural stem cells, transformed cells such as tumour cells, infected cells differentiated cells, etc. The cells may be targeted in culture or *in vivo*.

The vectors may be introduced into the cells by any convenient means. For example, with cells in culture, conventional techniques such as transfection (eg. lipofection, calcium phospate precipitation, electroporation, etc.), microinjection, viral infection etc. may be used. For cells within the organism, introduction may be mediated by virus, liposome, or any other convenient technique.

Examples of different trap vectors for use in this invention are as follows:

Gene trap: a promoterless reporter gene component carrying a splice acceptor sequence at the 5′ end. Each vector will splice in the correct reading frame in 1 of 3 possible reading frames. The vector needs to integrate in to an intron sequence of a gene creating a fusion transcript and a fusion protein product. The 5′

sequence of the fusion transcript is encoded by the endogenous gene and the 3′ sequence is encoded by the gene trap vector. Three gene trap vectors can be created for splicing into each of the three reading frames. The reporter gene lacks ATG translation start signal.

Gene trap (SA- IRES- reporter): a promoterless reporter gene containing a proper translation initiation sequence having a splice acceptor sequence (SA) followed by a IRES at the 5′ end.

Exon trap: a reporter gene component lacking a splice acceptor and a mammalian initiator codon. It needs to integrate into an exon coding sequence in the proper translational frame resulting in the production of a fusion protein.

Enhancer trap: reporter gene lacking a promoter but having a "minimal promoter" sequence encoding a TATA box and a transcriptional initiation site. These may be used to locate a target gene affected by an enhancer with restricted expression. The indicator must be monitored for enhanced expression in a particular cell type.

Promoter trap: promoterless reporter gene with appropriate translation initiation sequence. Insertion into an intron would not result in translation due to RNA processing. Integration immediately proximal to an endogenous promoter is required in order to have appropriate expression.

A splice acceptor (SA) includes the 3′ end of an intron and the 5′ end of an exon as described by Alberts, B. et al. at p. 373 of Molecular Biology of the Cell (1994), 3rd ed. Garland Publishing, N.Y. Splice Acceptors comprise a polypyrimidine tract followed by a nucleotide (N), followed by T or C and AG (3′ intron

portion), followed by at least G or A (5′ exon portion). Examples are the murine En-2 splice acceptor and the splice acceptors from human β-globin and rabbit b-globulin.

The following examples are offered by way of illustration of the invention and not by way of limitation.

Example 1

Complementation trap vectors may be developed based on the alpha complementation of beta-galactosidase.

Vectors

Figure 1 illustrates the CD11b alpha vector. The alpha subunit of beta-galactosidase is obtained by PCR amplification using the following oligos from the pCMV ß-galactosidase plasmid: oligo 1 corresponds to the polylinker sequence 5′ of β-gal ($^{5'}$AACTGCAGTACCCGCGGCCGC$^{3'}$ SEQ ID NO:1) and the oligo 2 introduces stop codons at a position corresponding to amino acid 85 of the E.coli lacZ sequence: ($^{5'}$AACTGCAGTTATTACTCAGGAAGATCGCACTCCAGC$^{3'}$ SEQ ID NO:2). The PCR fragment is subcloned into pBluescript™. The alpha subunit is then subcloned into the Bam HI site of the CD11b expression cassette containing the CD11b promoter and the human growth hormone mini gene and polyadenylation signal described in: Dzienni, S., et al. (1995) The CD11b Promoter Directs High-Level Expression of Reporter Genes in Macrophages in Transgenic Mice. Blood 85:319-329. The CD11b alpha vector also contains a pPGK Hygro pA cassette allowing for selection with hygromycin. The alpha fragment is also subcloned into the CMV promoter expression cassettes (Clontech).

The omega trap vector illustrated in **Figure 2**, is a derivative of pGT4.5 (Gossler et al, 1989 [supra]). The DNA fragment beginning at the unique BglII in the polylinker sequence between the En-2 splice acceptor sequence and the beginning of the β-gal sequence up to the unique ClaI site within the β-gal sequence from pFT4.5 is replaced by a PCR fragment amplified from pCMVbgal using the following oligos which delete the first 41aa from the N-terminal sequence of β-galactosidase resulting in the β-galactosidase sequence to begin at amino acid 42 of the E.coli β-galactosidase sequence to the unique ClaI site in that sequence. Oligo 1 ($^{5'}$CGGGATCCCAAGATCTGGGCCCGCACCGATCGCCCTTCC$^{3'}$ SEQ ID NO:3); Oligo 2 ($^{5'}$CGGGATCCATCGATAATTTCACCGCCG$^{3'}$ SEQ ID NO:4). PCR product is digested with BglII and ClaI and is directly inserted into the BglII-ClaI fragment of pGT4.5 vector. Three forms of this vector can be created; each resulting in a fusion in one of the three transcriptional reading frames.

Figure 3 illustrates the omega trap vector with the IRES sequence from EMCV (Clontech) inserted between the En-2 splice acceptor and the omega subunit. In this case, translation of the omega subunit will be independent of the endogenous translation initiation codon and start at the ATG present in the IRES sequence. This IRES sequence has been shown to work in all tissues and does not have any tissue specificity for IRES function. Other IRES suitable for use in vector of this invention are known and are derived from various viral sources including: EMCV, poliovirus, picornavirus, Hepatitis C virus, HTLV-1, Friend murine leukemia virus, Moloney murine leukemia virus, Theiler's murine encephalomyelitis virus, apthovirus, Hepatitis A virus, ECHO virus, rhinovirus, enterovirus, cardiovirus, and pestivirus.

Figure 4 illustrates an omega trap vector with a IRES sequence followed by a lox7 site between the omega gene and the poly

adenylation signal. This example may also be carried out using CD11b to drive the omega subunit and a trap vector comprising the alpha subunit.

The N terminal sequence of Neo to the unique Bal I site in Neo is fused to the omega fragment of β-galactosidase starting at the unique FspI site in β-galactosidase in the trap vectors illustrated in **Figures 2 - 4**. The fusion protein is driven by the Herpes thymidine kinase promoter and polyoma enhancers and has the SV40 polyA signal.

ß-gal Complementation

To test whether the omega fragment as a fusion protein is capable of complementation with the alpha fragment, vectors were constructed to express a portion of the Neo gene fused to the β-gal omega peptide described above under control of the human cytomegalovirus immediate - early gene (CMV) promoter. As a fusion with the neo gene product, the omega peptide was found to be capable of complementing with the alpha fragment in 293 kidney cell line. Transient transfection by Superfection™ (Qiagen) of 293 cells with alpha fragment driven by CMV promoter alone or the omega fragment alone showed background levels of enzyme activity whereas the combination of the alpha and omega subunits resulted in detectable enzyme activity by the luminescent β-galactosidase Genetic Reporter System II™ from Clontech.

ES Cell Culture, Selection and Gene Insertion

Murine R1 ES cells from A Nagy [Proc. Nat. Acad. Sci. U.S.A. (1993) 90, 8424-8428] may be grown on Primary Embryonic Fibroblast feeder layers or on gelatinized dishes in the presence of 1000 U/ml murine leukemia inhibitory factor (LIF), ESGRO™ (GIBCO BRL).

Selection conditions are: 150 μg/ml G418, 1.0 μg/ml puromycin, 110 μg/ml Hygromycin B. 2 x 10^7 R1 cells may be electroporated with 100μg linearized DNA in 0.8 ml PBS at 500μF and 240V with a BioRad Gene Pulser™ at room temperature.

The ES cells are electroported with CD11b promoter driving the alpha subunit and selected with hygromycin to obtain stable clones. Clones are selected and characterized by the expression pattern of the alpha subunit. Clones are isolated which express the alpha subunit in myeloid cells. The omega trap vector is electroported into CD11b alpha ES cells. Cells are plated into 96 well plates to obtain approximately 10 Neo resistant clones per well with 150 μg/ml G418. Clones are replica plated in duplicate. One set is frozen and one set is allowed to differentiate and assayed for β-gal activity. Positive wells are identified and individual cells subcloned and assay is repeated on individual clones. The following protocol may be used.

Picking of individual ES cell colonies:

1) plate out cells at low dilution;

2) grow for 9-12 days allowing individual cells to grow into individual colonies;

3) pick individual colonies in 5 to 10 μl volume using a P20 Gilson Pipetman™;

4) transfer colonies into 5 μl Trypsin/EDTA Gibco/BFL) in a 96 well round bottom plate;

5) incubate for 5 min at RT;

6) add 50 μl fresh serum containing ES cell media

7) disrupt colonies to single cell suspension by pipetting up and down; and

8) replate cells on gelatinized tissue culture plates and add fresh media as necessary.

Freezing ES cells in 96-well plates:

1) aspirate off media;

2) replace with 30 μl freezing media per well. (Freezing media contains: 50% serum, 10% DMSO, 40% DMEM, 1000U/ml LIF); and

3) transfer plate into -70°C freezer in a styrofoam box.

Thawing ES cells in 96 well plates:

1) add 100 μl prewarmed ES cell media per well;

2) incubate cells at 37°C 5% $C0_2$ for 6-8 hours; and

3) replace with fresh media.

Expression patterns of the alpha subunit may be characterized by *in situ* hybridization or immunohistochemistry using a monoclonal or anti-peptide antibody raised against the alpha peptide; immunofluorescence dual labelling with anti-alpha peptide and anti-MAC-1 (CD11b); or, by transfecting cells with CMV promoter driving the expression of the omega fragment and staining for β-gal activity.

In an alternative procedure, the ES cells are electroported with CD11b promoter driving expression of alpha subunit and selected with hygromycin to obtain stable clones. Several clones are identified and the expression pattern of the alpha subunit characterized. Clones are isolated which express the alpha subunit in myeloid cells. The omega trap vector containing IRES lox71 is electroplated into the CD11b alpha ES cell. The cells are plated in a 96 well plate to obtain approximately 10 Neo resistant clones per well. Clones are selected with 150 μg/ml G418 and replica plated in duplicate. One set is frozen and one set allowed to differentiate and assayed for β-gal activity. Positive wells are identified and individual cells subcloned and the assay is repeated on individual clones. Heterologous DNA is introduced into trapped events by co-electroporation of a transgene as a circular plasmid containing a single mutant lox (lox66) site with CRE expression cassette for transient expression of CRE. Stable integration of the transgene is determined by selection for a resistance gene found on transgene vector (eg. Araki et al., 1997 Targeted integration of DNA using mutant lox sites in embryonic stem cells. Nucleic Acids Research 25,868-872).

Preferential differentiation of ES cells to obtain macrophage cells may be accomplished by known procedures, including those described by: Lieschke, G.J. and Dunn, A.R. Development of Functional Macrophages from Embryonal Stem Cells *In Vitro*. (1995) Experimental Hematology 23(4):328-334. Alternatively, chimeric embryos and resulting chimeric mice may be obtained by injection of the ES cells into blastocysts, eg. C57BL/6 blastocysts.

5' RACE Cloning

5' Rapid amplification of cDNA ends (RACE) may be carried out as described by Skarnes, et al. at (1992) Genes and Development 6,

903-918, to clone a portion of the endogenous gene flanking the omega component insertion. This provides fragments for sequencing and to probe for genes. The source of reagents may be the 5' RACE kit from Gibco-BRL. An additional round of nested PCR may be performed. Macrophage cells may be enriched by FACsorting or immuno-magnetic bead purification prior to RNA isolation. PCR products may be subcloned into pBluescript™. In order to confirm the sequences of the trapped exons, 5 - 10 bacterial colonies should be picked from each line and the isolated plasmid DNAs sequenced by standard double strand sequencing protocols.

Example 2: Transcriptional Activation (Two hybrid system)

The two hybrid system is based on the fact that many eucaryotic transcriptional activators are comprised of two physically and functionally separable domains, a DNA-binding domain (DNA-BP) and an activation domain (AD). The two domains are normally part of the same protein. However, the two domains can be separated and expressed as distinct proteins. Two additional proteins (X and Y) are expressed as fusions to the DNA-BP and AD peptides. If X and Y interact, the AD is co-localized to the DNA-BP bound to the promoter, resulting in the transcription of the reporter gene.

The following is an example of the two hybrid transcriptional activation gene trap system. This system is composed of two fusion proteins, one component of which is expressed by a tissue- or cell-specific promoter and the second is found in a gene trap vector:

1) a fusion of protein X with the GAL4 DNA-BP;

2) a fusion of protein Y with the VP16 activation;

where protein X and Y interact (for example: the SV40 large T antigen which associates with the p53 protein). A third vector provides the GAL4 DNA binding site, the minimal promoter of the adenovirus E1b, and the lacZ reporter gene.

ES cells are first "seeded" with the third vector and a tissue-restricted promoter driving one component, for example: the CD11b promoter driving the expression of the GAL4 DNA-BP fused to p53 peptide sequence. The vector may also contain a selectable marker such as PGK neo. The cells are then transfected with the second component found in a gene trap vector system. The gene trap vector may contain a gene encoding a fusion protein comprising of the VP16 activation domain fused to the SV40 large T antigen preceded by a splice acceptor sequence and a IRES. The vector may also contain a selectable marker such as PGK-hygro. The ES cells are replica plated and some cells are allowed to differentiate and are then screened for β-gal activity.

Another example would be to use the following fusion proteins:

1) the GAL4 DNA binding domain fused to the Fos leucine zipper domain (DFosLZ), and

2) the VP16 activation domain fused to the Jun leucine zipper (AJunLZ); or

1) the GAL4 DNA binding domain fused to the Jun leucine zipper domain (DJunLZ), and

2) the VP16 activation domain fused to the Fos leucine zipper (AFosLZ).

The construction of these vectors have been previously described in Dang et al., (1991) Molecular and Cellular Biology, 11:954-962, and components to create the vectors of this system (except leucine zipper components) may be obtained from Clontech-Mammalian Matchmaker™ two hybrid assay kit.

An example of a reporter system whose expression is dependent on the presence of two interacting fusion proteins is the G5E1bCAT vector which contains 5 copies of the 17 mer GAL4 DNA binding site 5′ of the minimal promoter of the Adenovirus E1b driving the expression of the CAT reporter gene. The CAT reporter gene may also be replaced by the β-gal reporter gene or any other reporter gene. In order to adapt this to the complementation gene trap system, one component for example DFosLZ is driven by a tissue restricted promoter for example the CD11b promoter. The DFosLZ gene can be subcloned into a CD11b expression cassette (eg. as described in Dziennis, S. et al. (1995), Blood 85:319-329). The second component will be contained in a gene trap vector system. The trap vector will be a variation of the IRES containing gene trap system described by Chowdhury et al., [*supra*] where the βgeo gene is replaced with the AJunLZ gene.

Example 3: Transcriptional Activation (Genetic Recombination Dependent)

This system makes use of a vector containing a stuffer DNA fragment flanked by lox sites preceding a reporter gene similar to that described by Lasko, et al., (1992): Targeted Oncogene Activation by Site-Specific Recombination in Transgenic Mice; Proc. Natl. Acad. Sci. USA 89:6232-6236. The expression of the reporter gene is dependent on the removal of the stuffer DNA sequence mediated by the Cre recombinase enzyme through site directed recombination of lox sites found flanking the stuffer DNA. The gene trap vector contains the Cre gene preceded by a splice

acceptor sequence and an IRES. The gene trap vector may also contain a selectable marker.

Example 4: Selection/Screening System

A selectable marker is driven by a tissue restricted promoter. For example, the myeloid specific promoter CD11b driving the expression of the puromycin resistance gene (puromycin n-acetyl transferase, *pac*). The vector may also contain PGK-hygro for ES cell selection. The cells are then transfected with a standard gene trap vector such as pGT4.5 which contains the lacZ reporter, and then allowed to differentiate and subsequently placed under selection with puromycin. Only myeloid specific cells survive the selection protocol. The cells are then assayed for the expression of the reporter gene, β-gal.

Example 5: Co-expression of Two Reporter Genes

An example of this is where ES cells are first seeded with one reported such as the human CD8 cell surface protein driven by a myeloid specific promoter such as CD11b. The vector may also contain a selectable marker such as PGK hygro for ES cell clone selection. The cells are then transfected with a gene trap vector comprising the green fluorescent protein preceded by a splice acceptor sequence and an IRES sequence. The gene trap vector may also contain a selectable marker such as PGK neo. Individual clones are replica plated and some cells are allowed to undergo differentiation and are subsequently screened for dual expression of the reporter genes by immunofluorescence and FACS analysis. The cells are stained with a monoclonal antibody directed against the human CD8 cell surface protein that is conjugated to phyco-erythrin. Cells are then subsequently analysed for red and green fluorescence by Fluorescent Activated Cell Sorting analyses.

All publications and patents cited in this specification are incorporated herein by reference. Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in the art in light of the teachings of this invention that changes and modification may be made thereto without departing from the spirit or scope of the appended claims.

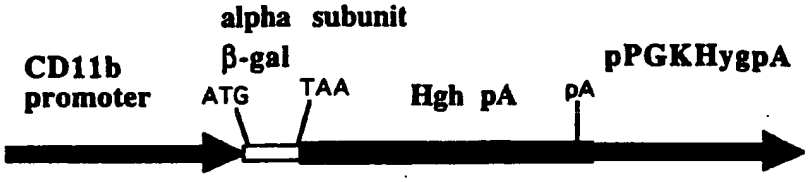

FIGURE 1

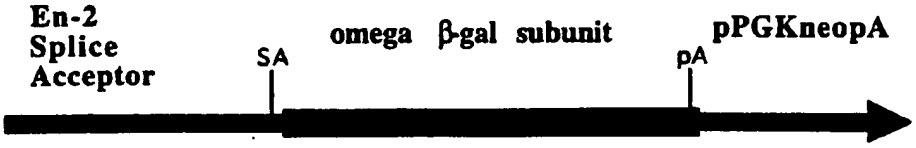

FIGURE 2

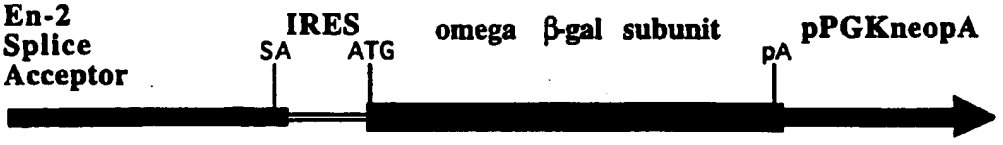

FIGURE 3

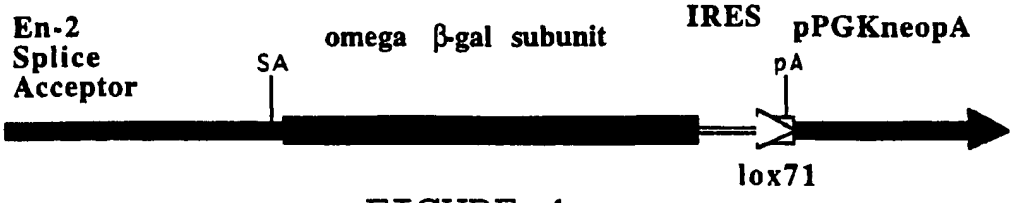

FIGURE 4